REMARKS

Claim 21 was objected to because it depended from canceled claim 20. With the present amendment, claim 21 has been amended to depend from claim 19.

Claims 1-3, 5-8, 10, 12-19, 21 and 22 have been rejected either under 35 U.S.C. §102(b) as being anticipated by Su et al. (U.S. Patent 5,418,717, hereinafter Su), or under 35 U.S.C. §103(a) as being unpatentable over Su in view of Kucera et al. (U.S. Patent 4,868,750, hereinafter Kucera).

Independent claims 1, 10 and 19 each include a limitation to a mutual information score that is based in part on a phrase level. Neither Su nor Kucera show or suggest a mutual information score that is based on a phrase level. In the specification of the present application, mutual information is defined as measuring the correlation between two or more events. Neither Kucera nor Su provide a mutual information score that provides the correlation between two or more events.

In the Final Office Action, it was asserted that Su suggests employing a mutual information metric as part of the parsing process in that it uses a conditional probability for determining its syntactic score. However, a conditional probability is not the same as mutual information. In particular, a conditional probability does not provide a correlation between two or more events. It simply provides the probability of a first event given a second event. The distinction is that the conditional probability does not indicate how the first event influences the second event. For example, if a conditional probability of seeing a child with freckles given that you see a child with red hair is 0.5 there is no indication of how children with freckles relate to children with red hair because there is no information about the probability of seeing a child with freckles who does not have red hair. Thus, if the

probability of seeing any child with freckles in the classroom is 0.5, there would be no correlation between red hair and freckles. However, if the probability of seeing a child with freckles is 0.1, it would appear that red hair and freckles are strongly correlated since you are much more likely to see freckles in children who have red hair then in the general population. Thus, the conditional probability alone does not provide information about the correlation between two events. As such, the conditional probability shown in Su is not a mutual information metric.

The Final Office Action also asserted that Kucera shows a mutual information metric at column 2, lines 20-34 when it discusses its collocation probability. However, the collocation probability in Kucera is not a mutual information metric because it does not indicated the correlation between two events. This can be seen simply by thinking of a set of training data. If in the set of training data, two events occurred next to each other ten times and the individual events each occurred ten times, the collocation function in Kucera would evaluate to 10/(10x10) or 1/10. If we doubled the size of the training data simply by duplicating the set of training data, the score would then become 20/(20x20) or 1/20. However, since the same training data is found in both examples, the correlation between the two events should be the same in both sets of training data. The fact that the collocation function in Kucera changes in this example, indicates that it does not show the correlation between two events but instead is providing some other type of information. As such, the function shown in Kucera is not a mutual information metric.

In the Final Office Action, the Examiner pointed to the background of Lu (U.S. Patent No. 5,819,260), which makes a statement that "[t]he statistics used include collocation information or mutual information, i.e., the probability that a

given pair of a part-of-speech tags or a given pair of words tends to appear together in a given data collection." Lu goes on to site Kucera directly later in the same paragraph. However, Lu is not saying that Kucera uses mutual information. Further, Lu is not saying that collocation information is the same as mutual information. This is particularly true in Kucera where the collocation determination made by Kucera is clearly not a mutual information metric.

Since neither Su nor Kucera show a mutual information metric, they cannot show or suggest the inventions of claims 1, 10 or 19 which include using mutual information that is based on a phrase level. As such, the combination of Kucera and Su do not show or suggest the inventions of claims 1-3, 5-8, 10, 12-19, 21 and 22.

The Final Office Action also disputed Applicant's argument that claims 7, 13 and 21 were further patentable over Su and Kucera because in these claims, the mutual information score is further based on all possible word classes of a word in the text segment. In particular, the Examiner cited column 17, lines 47-66, column 11, lines 46-50 and FIG. 7 of Su as showing this limitation. However, as indicated by the Examiner, Su is not generating a single score based on all possible word classes of a word, but instead is calculating separate probabilities for each part of speech. The separate scores are then compared to each other to determine which sequence of parts-of-speech to use. No single score is formed from all of the word classes. This is substantially different from claims 7, 13 and 21, where a mutual information score is determined from all possible word classes.

In light of the fact that neither Kucera nor Su show or suggest generating a mutual information score based on a phrase level, all of the pending claims are in form for allowance. Reconsideration and allowance of the claims is respectfully requested.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____
     Theodore M. Magee, Reg. No. 39,758
     Suite 1400 - International Centre
     900 Second Avenue South
     Minneapolis, Minnesota 55402-3319
     Phone: (612) 334-3222 Fax: (612) 339-3312

TMM:sew